# RAW SEQUENCE LISTING

The Biotechnology Systems Branch of the Scientific and Technical Information Center (STIC) no errors detected.

Application Serial Number: _10/501,930 A_
Source: _1Fw/6_
Date Processed by STIC: _12/12/06_

# ENTERED

```
       RAW SEQUENCE LISTING                    DATE: 12/12/2006
       PATENT APPLICATION:  US/10/501,930A     TIME: 10:34:57

       Input Set : A:\P25687.APP
       Output Set: N:\CRF4\12122006\J501930A.raw

  3 <110> APPLICANT: TAKASHIMA, SHOU
  4        TSUJIMOTO, MASAFUMI
  5        TSUJI, SHUICHI
  7 <120> TITLE OF INVENTION: GLYCOSYLATING ENZYME
  9 <130> FILE REFERENCE: P25687
 11 <140> CURRENT APPLICATION NUMBER: 10/501,930A
C--> 12 <141> CURRENT FILING DATE: 2005-03-17
 14 <150> PRIOR APPLICATION NUMBER: PCT/JP03/00883
 15 <151> PRIOR FILING DATE: 2003-01-30
 17 <150> PRIOR APPLICATION NUMBER: JP 2002-21159
 18 <151> PRIOR FILING DATE: 2002-01-30
 20 <150> PRIOR APPLICATION NUMBER: JP 2002-122673
 21 <151> PRIOR FILING DATE: 2002-04-24
 23 <160> NUMBER OF SEQ ID NOS: 36
 25 <170> SOFTWARE: PatentIn Ver. 3.3
 27 <210> SEQ ID NO: 1
 28 <211> LENGTH: 398
 29 <212> TYPE: PRT
 30 <213> ORGANISM: Mus sp.
 32 <400> SEQUENCE: 1
 33 Met Arg Ser Gly Gly Thr Leu Phe Ala Leu Ile Gly Ser Leu Met Leu
 34  1               5                  10                  15
 36 Leu Leu Leu Leu Arg Met Leu Trp Cys Pro Ala Asp Ala Pro Ala Arg
 37              20                  25                  30
 39 Ser Arg Leu Leu Met Glu Gly Ser Arg Glu Asp Thr Ser Gly Thr Ser
 40          35                  40                  45
 42 Ala Ala Leu Lys Thr Leu Trp Ser Pro Thr Thr Pro Val Pro Arg Thr
 43      50                  55                  60
 45 Arg Asn Ser Thr Tyr Leu Asp Glu Lys Thr Thr Gln Ile Thr Glu Lys
 46  65                  70                  75                  80
 48 Cys Lys Asp Leu Gln Tyr Ser Leu Asn Ser Leu Ser Asn Lys Thr Arg
 49              85                  90                  95
 51 Arg Tyr Ser Glu Asp Asp Tyr Leu Gln Thr Ile Thr Asn Ile Gln Arg
 52              100                 105                 110
 54 Cys Pro Trp Asn Arg Gln Ala Glu Glu Tyr Asp Asn Phe Arg Ala Lys
 55          115                 120                 125
 57 Leu Ala Ser Cys Cys Asp Ala Ile Gln Asp Phe Val Val Ser Gln Asn
 58      130                 135                 140
 60 Asn Thr Pro Val Gly Thr Asn Met Ser Tyr Glu Val Glu Ser Lys Lys
 61 145                 150                 155                 160
 63 His Ile Pro Ile Arg Glu Asn Ile Phe His Met Phe Pro Val Ser Gln
 64              165                 170                 175
 66 Pro Phe Val Asp Tyr Pro Tyr Asn Gln Cys Ala Val Val Gly Asn Gly
```

```
67              180                 185                 190
69 Gly Ile Leu Asn Lys Ser Leu Cys Gly Ala Glu Ile Asp Lys Ser Asp
70            195                 200                 205
72 Phe Val Phe Arg Cys Asn Leu Pro Pro Ile Thr Gly Ser Ala Ser Lys
73      210                 215                 220
75 Asp Val Gly Ser Lys Thr Asn Leu Val Thr Val Asn Pro Ser Ile Ile
76 225                 230                 235                 240
78 Thr Leu Lys Tyr Gln Asn Leu Lys Glu Lys Lys Ala Gln Phe Leu Glu
79            245                 250                 255
81 Asp Ile Ser Thr Tyr Gly Asp Ala Phe Leu Leu Leu Pro Ala Phe Ser
82            260                 265                 270
84 Tyr Arg Ala Asn Thr Gly Ile Ser Phe Lys Val Tyr Gln Thr Leu Lys
85            275                 280                 285
87 Glu Ser Lys Met Arg Gln Lys Val Leu Phe Phe His Pro Arg Tyr Leu
88      290                 295                 300
90 Arg His Leu Ala Leu Phe Trp Arg Thr Lys Gly Val Thr Ala Tyr Arg
91 305                 310                 315                 320
93 Leu Ser Thr Gly Leu Met Ile Ala Ser Val Ala Val Glu Leu Cys Glu
94            325                 330                 335
96 Asn Val Lys Leu Tyr Gly Phe Trp Pro Phe Ser Lys Thr Ile Glu Asp
97            340                 345                 350
99 Thr Pro Leu Ser His His Tyr Tyr Asp Asn Met Leu Pro Lys His Gly
100           355                 360                 365
102 Phe His Gln Met Pro Lys Glu Tyr Ser Gln Met Leu Gln Leu His Met
103      370                 375                 380
105 Arg Gly Ile Leu Lys Leu Gln Phe Ser Lys Cys Glu Thr Ala
106 385                 390                 395
109 <210> SEQ ID NO: 2
110 <211> LENGTH: 3166
111 <212> TYPE: DNA
112 <213> ORGANISM: Mus sp.
114 <220> FEATURE:
115 <221> NAME/KEY: CDS
116 <222> LOCATION: (77)..(1270)
118 <400> SEQUENCE: 2
119 cggagcggcg agtcggtgcc gcccgggctg cgcttcgccc cggcagcttt ggcggcgagg 60
121 acgcccgtgg ctcagg atg aga tcg ggg ggc acg ctg ttc gcc ctc ata ggc 112
122                Met Arg Ser Gly Gly Thr Leu Phe Ala Leu Ile Gly
123                 1               5                   10
125 agc ctg atg ctg ctg ctc ctc ctg cgt atg ctc tgg tgc cca gcc gac    160
126 Ser Leu Met Leu Leu Leu Leu Leu Arg Met Leu Trp Cys Pro Ala Asp
127       15                 20                  25
129 gcg cct gcc cgc tcc agg ctg ttg atg gag gga agc aga gag gac acc    208
130 Ala Pro Ala Arg Ser Arg Leu Leu Met Glu Gly Ser Arg Glu Asp Thr
131      30                 35                  40
133 agt ggt acc tca gct gca ctg aag aca ctc tgg agc ccg aca acc ccg    256
134 Ser Gly Thr Ser Ala Ala Leu Lys Thr Leu Trp Ser Pro Thr Thr Pro
135 45                 50                  55                  60
137 gta cca cgc acc agg aac agc aca tat ctg gat gag aag aca acc caa    304
```

RAW SEQUENCE LISTING                          DATE: 12/12/2006
PATENT APPLICATION:   US/10/501,930A          TIME: 10:34:57

Input Set : A:\P25687.APP
Output Set: N:\CRF4\12122006\J501930A.raw

```
138 Val Pro Arg Thr Arg Asn Ser Thr Tyr Leu Asp Glu Lys Thr Thr Gln
139             65                  70                  75
141 ata aca gag aaa tgc aaa gat ctg caa tat agc ttg aac tct tta tct    352
142 Ile Thr Glu Lys Cys Lys Asp Leu Gln Tyr Ser Leu Asn Ser Leu Ser
143             80                  85                  90
145 aac aaa acg aga cgg tac tct gag gat gac tac ctc cag acc atc aca    400
146 Asn Lys Thr Arg Arg Tyr Ser Glu Asp Asp Tyr Leu Gln Thr Ile Thr
147             95                 100                 105
149 aac ata cag aga tgc cca tgg aac cgg caa gca gaa gaa tat gac aat    448
150 Asn Ile Gln Arg Cys Pro Trp Asn Arg Gln Ala Glu Glu Tyr Asp Asn
151         110                 115                 120
153 ttt aga gca aaa ctg gct tcc tgt tgc gat gcc att caa gac ttc gtg    496
154 Phe Arg Ala Lys Leu Ala Ser Cys Cys Asp Ala Ile Gln Asp Phe Val
155 125                 130                 135                 140
157 gtt tcc cag aac aac act cca gtg ggg act aac atg agc tac gag gtg    544
158 Val Ser Gln Asn Asn Thr Pro Val Gly Thr Asn Met Ser Tyr Glu Val
159             145                 150                 155
161 gaa agc aag aaa cac atc ccc att cga gag aac att ttc cac atg ttt    592
162 Glu Ser Lys Lys His Ile Pro Ile Arg Glu Asn Ile Phe His Met Phe
163             160                 165                 170
165 cca gtg tcg cag cct ttt gtg gac tat ccc tat aac cag tgt gca gtg    640
166 Pro Val Ser Gln Pro Phe Val Asp Tyr Pro Tyr Asn Gln Cys Ala Val
167             175                 180                 185
169 gtt ggt aat ggg gga att ctc aac aag tct ctc tgc gga gca gaa att    688
170 Val Gly Asn Gly Gly Ile Leu Asn Lys Ser Leu Cys Gly Ala Glu Ile
171         190                 195                 200
173 gat aaa tct gac ttc gtc ttc agg tgt aac ctc ccc cca atc aca ggg    736
174 Asp Lys Ser Asp Phe Val Phe Arg Cys Asn Leu Pro Pro Ile Thr Gly
175 205                 210                 215                 220
177 agc gct agt aaa gat gtt gga agc aaa aca aat ctt gtg act gtc aat    784
178 Ser Ala Ser Lys Asp Val Gly Ser Lys Thr Asn Leu Val Thr Val Asn
179             225                 230                 235
181 ccc agc att ata acc ctg aag tac cag aat ttg aag gag aag aaa gca    832
182 Pro Ser Ile Ile Thr Leu Lys Tyr Gln Asn Leu Lys Glu Lys Lys Ala
183         240                 245                 250
185 cag ttt ttg gag gac atc tcc acc tat gga gat gca ttc ctc ctc ctg    880
186 Gln Phe Leu Glu Asp Ile Ser Thr Tyr Gly Asp Ala Phe Leu Leu Leu
187         255                 260                 265
189 cca gca ttt tcc tat cgg gcc aac aca ggc atc tct ttt aaa gtc tac    928
190 Pro Ala Phe Ser Tyr Arg Ala Asn Thr Gly Ile Ser Phe Lys Val Tyr
191     270                 275                 280
193 caa aca ctc aaa gag tca aaa atg agg caa aag gtt ctc ttc ttc cat    976
194 Gln Thr Leu Lys Glu Ser Lys Met Arg Gln Lys Val Leu Phe Phe His
195 285                 290                 295                 300
197 ccc agg tac ctg aga cac ctc gct ctt ttc tgg aga act aaa ggg gtg    1024
198 Pro Arg Tyr Leu Arg His Leu Ala Leu Phe Trp Arg Thr Lys Gly Val
199             305                 310                 315
201 act gca tac cgc ttg tcc aca ggc ttg atg att gca agt gtc gct gtg    1072
202 Thr Ala Tyr Arg Leu Ser Thr Gly Leu Met Ile Ala Ser Val Ala Val
```

```
                    RAW SEQUENCE LISTING              DATE: 12/12/2006
                    PATENT APPLICATION:  US/10/501,930A    TIME: 10:34:57

                    Input Set : A:\P25687.APP
                    Output Set: N:\CRF4\12122006\J501930A.raw


203              320                 325                 330
205 gaa ctg tgt gaa aac gtg aag ctc tac gga ttc tgg cct ttc tct aag   1120
206 Glu Leu Cys Glu Asn Val Lys Leu Tyr Gly Phe Trp Pro Phe Ser Lys
207          335                 340                 345
209 act atc gaa gac acc cca ctc agt cac cac tac tat gat aac atg tta   1168
210 Thr Ile Glu Asp Thr Pro Leu Ser His His Tyr Tyr Asp Asn Met Leu
211      350                 355                 360
213 cct aag cat ggt ttc cac cag atg cct aaa gaa tac agc caa atg ctc   1216
214 Pro Lys His Gly Phe His Gln Met Pro Lys Glu Tyr Ser Gln Met Leu
215 365                 370                 375                 380
217 cag ctc cat atg aga gga atc ctc aaa ctg caa ttc agc aaa tgt gaa   1264
218 Gln Leu His Met Arg Gly Ile Leu Lys Leu Gln Phe Ser Lys Cys Glu
219              385                 390                 395
221 acg gct taacgtttct tagaaggaga ataatttcag gaggtggagt ggatgtgtca   1320
222 Thr Ala
224 cagcatctcc aaaaagccaa tagaagaagg cacagagaaa gcatgaatta caaaggcgct   1380
226 ctcccacttg tctagaccaa agccacccgc ccccactcac tttgcagcct ccacgagtca   1440
228 ctcattctca ccttcaacgt tctttctctg agaatagaga ccaaaacatc agacttggat   1500
230 aagtaaaatg agataatttt tcaaatcatc atagaatttg atttgagcca gggtctctca   1560
232 gaatgcttcc ttgttcctat ccatgatagc cattcccacc tttatcagag tggtaatgaa   1620
234 actgtgcaat tgtgccaaag accctttctg aagagaatgt ctgaatcatg cgccgagttt   1680
236 ttacacacag ctcttccttt ataaataaat ccttcccatt ctccctccta gtagagtaca   1740
238 gaaacaaaat acccttgatg attcaggaag aaaagtcttt tttacttagc aatgtgcctg   1800
240 cttctgattc agttcgcttg tgacattaag ctgggttggg gttttggttg gatttggggc   1860
242 gtttcttcac ttcttttgtc tatattttcc ttacctttat cagtttgtat tcgagcttcc   1920
244 tgctttggga ttctgcaatt ctctctccca ctgacaggat caactcaatg acataaagta   1980
246 gttcaaacat ccattgcttc tcacatgttt tatccataaa gttactcatc tgattttatt   2040
248 taaaatagtg aacatctact tgatatcaga cccgaggacc atcctccatt ggagaatatg   2100
250 aagatattgt cactggcaga aaagcaggtg tgtgccatta attgataaga taccacaagc   2160
252 atcatcatgc cagttatgaa cacagtgctg aaaggatcat agacaggggt ggttaaatct   2220
254 gatcccagta gaataaactt cagtgtacct atttcaggga agagttaatt tcacaattaa   2280
256 aactagtaaa tgaaccaatt cttaggcaca ttaagtggat tctgagtaaa agaaagggaa   2340
258 cagcaggaga aagctgttcg cttggttctg attacccaaa tgagcatgct ggaaggaggt   2400
260 tgtgaggcta cgctaaaacc tctgcgtagg gagagagtac agtgcatgag tgtggcggct   2460
262 tttgtccaca ctcgtgaagg gtgagtaatt cagagccaat cacatcacaa ggatggacac   2520
264 acctaactca tcacttcagg gggagatgaa tgctttcatg agaaattaca ctcataagct   2580
266 aagcatcagt tttgagtaaa atttgagtag atgttaaata tgaacatttt atacctctta   2640
268 ctaatgtccc accgacacct tttaatgtaa gcacatttat ttattaagtt acttgacatt   2700
270 aaatgcttat gtctgtatat tctgttcatc catcgatttt cccaaaaagt aagagcatag   2760
272 gagatgaggc ctacatgcca agaaaactat aaattttact ctttaattct tacttgagcc   2820
274 agcttgttgt ttatcaagtg cttttttgaa gagacagcac cctgtgaatt cttcattctg   2880
276 atacagtgtc accttgtatt taacatttgt aatgttgttt caagtttaca tctctttcat   2940
278 tcttttatag caaatcaaac gtattagctt cagaaattta tcagaagttc atatataaat   3000
280 attttgcaaa gggtaaaagg cttttttgtt aaataaaata aaatttatta ttttcttctg   3060
282 atgaatagag gctcttttat gctgctgcta atgaacctaa ttagctttaa attatctcct   3120
284 agcaacattg gtcacgtttc aatcatgcta ttagcaaaaa aaaaaa               3166
287 <210> SEQ ID NO: 3
288 <211> LENGTH: 398
289 <212> TYPE: PRT
```

```
290 <213> ORGANISM: Homo sapiens
292 <400> SEQUENCE: 3
293 Met Arg Pro Gly Gly Ala Leu Leu Ala Leu Leu Ala Ser Leu Leu Leu
294  1                5                  10                      15
296 Leu Leu Leu Leu Arg Leu Leu Trp Cys Pro Ala Asp Ala Pro Gly Arg
297              20                  25                  30
299 Ala Arg Ile Leu Val Glu Glu Ser Arg Glu Ala Thr His Gly Thr Pro
300          35                  40                  45
302 Ala Ala Leu Arg Thr Leu Arg Ser Pro Ala Thr Ala Val Pro Arg Ala
303      50                  55                  60
305 Thr Asn Ser Thr Tyr Leu Asn Glu Lys Ser Leu Gln Leu Thr Glu Lys
306  65                  70                  75                  80
308 Cys Lys Asn Leu Gln Tyr Gly Ile Glu Ser Phe Ser Asn Lys Thr Lys
309              85                  90                  95
311 Gly Tyr Ser Glu Asn Asp Tyr Leu Gln Ile Ile Thr Asp Ile Gln Ser
312          100                 105                 110
314 Cys Pro Trp Lys Arg Gln Ala Glu Glu Tyr Ala Asn Phe Arg Ala Lys
315          115                 120                 125
317 Leu Ala Ser Cys Cys Asp Ala Val Gln Asn Phe Val Val Ser Gln Asn
318      130                 135                 140
320 Asn Thr Pro Val Gly Thr Asn Met Ser Tyr Glu Val Glu Ser Lys Lys
321 145                 150                 155                 160
323 Glu Ile Pro Ile Lys Lys Asn Ile Phe His Met Phe Pro Val Ser Gln
324              165                 170                 175
326 Pro Phe Val Asp Tyr Pro Tyr Asn Gln Cys Ala Val Val Gly Asn Gly
327              180                 185                 190
329 Gly Ile Leu Asn Lys Ser Leu Cys Gly Thr Glu Ile Asp Lys Ser Asp
330          195                 200                 205
332 Phe Val Phe Arg Cys Asn Leu Pro Pro Thr Thr Gly Asp Val Ser Lys
333      210                 215                 220
335 Asp Val Gly Ser Lys Thr Asn Leu Val Thr Ile Asn Pro Ser Ile Ile
336 225                 230                 235                 240
338 Thr Leu Lys Tyr Gly Asn Leu Lys Glu Lys Lys Ala Leu Phe Leu Glu
339              245                 250                 255
341 Asp Ile Ala Thr Tyr Gly Asp Ala Phe Phe Phe Leu Pro Ala Phe Ser
342          260                 265                 270
344 Phe Arg Ala Asn Thr Gly Thr Ser Phe Lys Val Tyr Tyr Thr Leu Glu
345          275                 280                 285
347 Glu Ser Lys Ala Arg Gln Lys Val Leu Phe Phe His Pro Lys Tyr Leu
348      290                 295                 300
350 Lys Asp Leu Ala Leu Phe Trp Arg Thr Lys Gly Val Thr Ala Tyr Arg
351 305                 310                 315                 320
353 Leu Ser Thr Gly Leu Met Ile Thr Ser Val Ala Val Glu Leu Cys Lys
354              325                 330                 335
356 Asn Val Lys Leu Tyr Gly Phe Trp Pro Phe Ser Lys Thr Val Glu Asp
357              340                 345                 350
359 Ile Pro Val Ser His His Tyr Tyr Asp Asn Lys Leu Pro Lys His Gly
360              355                 360                 365
362 Phe His Gln Met Pro Lys Glu Tyr Ser Gln Ile Leu Gln Leu His Met
```

**VERIFICATION SUMMARY**          DATE: 12/12/2006
PATENT APPLICATION:  **US/10/501,930A**     TIME: 10:34:58

Input Set : **A:\P25687.APP**
Output Set: **N:\CRF4\12122006\J501930A.raw**

L:12 M:271 C: Current Filing Date differs, Replaced Current Filing Date